

University of Groningen

Quantification of symptoms of movement disorders employing motion sensors

Martinez Manzanera, Octavio

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Martinez Manzanera, O. (2017). *Quantification of symptoms of movement disorders employing motion sensors*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

NEDERLANDSE SAMENVATTING

Inleiding

Bewegingsstoornissen zijn neurologische aandoeningen die gekenmerkt worden door ofwel een overmaat aan bewegingen (hyperkinesie) ofwel een afname van doelgerichte en automatische bewegingen (hypokinesie), waarbij deze bewegingsvermindering niet het gevolg is van vermoeidheid [1]. Bewegingsstoornissen hebben verschillende mogelijke oorzaken. De symptomen van een patiënt hangen af van de ernst van de bewegingsstoornis en van het hersengebied dat is aangedaan. Ondanks deze variatie overlappen de symptomen van veel bewegingsstoornissen. Zo komt een pathologische tremor voor bij patiënten met de ziekte van Parkinson (ZvP, engels: PD), de belangrijkste oorzaak van invaliditeit bij ouderen, maar ook bij mensen met een essentiële tremor (ET), dat één van de meest voorkomende bewegingsstoornissen is [2]. Een ander voorbeeld is de overeenkomst tussen ‘early-onset ataxia’ (EOA) en ‘developmental coordination disorder’ (DCD; ook bekend als dyspraxie), die allebei worden gekenmerkt door verstoorde coördinatie.

Het diagnostisch proces bij patiënten met bewegingsstoornissen bestaat uit verschillende onderdelen. Eerst wordt een anamnese afgenomen, vervolgens wordt de patiënt uitvoerig onderzocht en ten slotte kan, indien noodzakelijk, beeldvormend onderzoek of laboratoriumdiagnostiek worden aangevraagd. Het belangrijkste diagnostische onderdeel is het lichamelijk onderzoek. Hierbij probeert de arts de symptomen van de bewegingsstoornis nauwkeurig in kaart te brengen om daarmee aanwijzingen voor de ziekteoorzaak te verzamelen. Dat is niet altijd eenvoudig, omdat de symptomen van bewegingsstoornissen vaak overlappen. Het vereist dan ook expertise en veel klinische ervaring om bij een patiënt met een niet-eenduidige bewegingsstoornis de juiste diagnose te stellen. Expertise en ervaring zijn ook belangrijk bij het beoordelen of een aandoening in ernst toeneemt of afneemt, en bij het inschatten of een behandeling effectief is. Ook hiervoor geldt dat het moeilijk is om subtiele veranderingen waar te nemen.

Vanuit de behoefte om de beoordeling van bewegingsstoornissen te standaardiseren, zijn er verschillende klinimetrische schalen ontwikkeld. De ‘Unified Parkinson’s Disease Rating Scale’ (UPDRS) dateert uit de jaren ‘80 en is de meest gebruikte schaal bij patiënten met de ziekte van Parkinson. Ook voor andere aandoeningen zijn schalen ontwikkeld, zoals de ‘Scale for the Assessment and Rating of Ataxia’ (SARA) die gebruikt wordt bij patiënten met ataxie [3]. Daarnaast zijn er schalen die één specifiek symptoom van een bewegingsstoornis kunnen meten, zoals de ‘Modified Bradykinesia Rating Scale’

(MBRS) voor bradykinesie [4]. De meeste schalen bestaan uit een reeks opdrachten die de patiënt moet uitvoeren en instructies voor de beoordelaar hoe opdrachten moeten worden gescord. Hoewel er is geprobeerd om de schalen en het gebruik ervan zoveel mogelijk te standaardiseren, zal de beoordeling met behulp van zulke schalen altijd enigszins subjectief blijven.

In dit proefschrift worden verschillende onderzoeken naar de verbetering van zowel de diagnostiek als monitoring van patiënten met bewegingsstoornissen besproken. Iedere studie had als doel om met bewegingssensoren specifieke symptomen uit te drukken in maat en getal. De verzamelde meetgegevens werden vervolgens gebruikt om bewegingsstoornissen van elkaar te onderscheiden of om uitspraken te doen over de ernst van de symptomen.

Methode en Resultaten

Bewegingssensoren meten een specifiek fysiek fenomeen dat kan worden gerelateerd aan de beweging van een object. Er zijn verschillende typen sensoren die kunnen worden toegepast voor het meten van verschillende bewegingseigenschappen. Accelerometers, bijvoorbeeld, worden gebruikt voor het meten van lineaire acceleratie, en gyroscopen meten de hoeksnelheid. Beide typen sensoren kunnen worden gebruikt voor het meten van bewegingen van het menselijk lichaam. De signalen die gemeten worden door de sensoren zijn echter verschillend. Dit geldt ook als ze worden onderworpen aan dezelfde fysieke omstandigheden en op dezelfde positie worden geplaatst. Verder kan iedere sensor, afhankelijk van de werkingsprincipes, gevoeliger zijn voor bepaalde soorten ruis of beïnvloed worden door specifieke artefacten. Daarom is het belangrijk om te weten dat accelerometers, naast het meten van de acceleratie van een object waaraan ze gekoppeld zijn, ook in staat zijn om acceleratie als gevolg van statische krachten, zoals zwaartekracht, te meten. Gyroscopen kunnen weer worden beïnvloed door andere typen artefacten, zoals drift. Deze beperkingen zijn van belang voor het begrijpen van de toepassing van sensoren. De combinatie van data van twee of meer sensoren wordt sensorfusie genoemd en wordt toegepast om een betere schatting van de onderzochte variabele, bijvoorbeeld acceleratie of snelheid, te verkrijgen. In dit proefschrift is sensorfusie gebruikt om gegevens te combineren van accelerometers, gyroscopen en magnetometers. Dit wordt toegelicht in hoofdstukken III, IV en V. De gegevens van de accelerometer en de gyroscopen zijn geanalyseerd in, respectievelijk, hoofdstuk II en hoofdstuk VI.

Overzicht hoofdstukken

In **Hoofdstuk II** worden verschillende automatische tremor detectiemethoden toegepast op accelerometer data en vergeleken, met als doel om hiermee de meest geschikte methode voor het detecteren van tremor in langdurende metingen te bepalen. Tremor is het meest voorkomende bewegingsstoornis symptoom bij ouderen en is aanwezig bij verschillende bewegingsstoornissen zoals PD, ET en functionele tremor. De karakteristieken van tremor kunnen helpen bij het identificeren van de geassocieerde bewegingsstoornis. De eerste stap is daarom het vastleggen van bewegingsdata waarin tremor aanwezig is. Hierna kan een ervaren beoordelaar perioden van tremor identificeren door middel van het detecteren van specifieke patronen in de gemeten data. Op deze manier werden de tremor data van 14 patiënten met verschillende bewegingsstoornissen geanalyseerd. De tremor data werden verkregen met behulp van accelerometers op de handrug van de patiënten, terwijl zij werden onderworpen aan een routine evaluatie in de kliniek. Het vermogen van iedere methode om tremor te herkennen werd geëvalueerd met behulp van de F_1 score die zowel precisie als recall, ofwel sensitiviteit (gelijk aan recall) en specificiteit, integreert. Alle methoden behaalden een gelijke prestatie in termen van de F_1 score. Over het algemeen presteerden niet-parametrische methoden (op basis van de Fourier transformatie) beter dan parametrische methoden (op basis van autoregressie). Interessant is dat één parametrisch methode ('hoge frequentie') de hoogste recall behaalde ten opzichte van alle andere methoden.

Het doel van **Hoofdstuk III** was het ontwikkelen van een automatische classificatie methode van de ernst van bradykinesie. Bradykinesie is gedefinieerd als vertraagde beweging en is een belangrijk symptoom van PD. Een Inertial Measurement Unit of IMU werd gebruikt om bewegingen te meten van 25 patiënten met PD en 10 controlepersonen, terwijl zij drie van de vijf bradykinesie gerelateerde items, zoals gedefinieerd in de UPDRS (tikken met de vinger, diadochokinese en tikken met de tenen), uitvoerden. De IMU bestaat uit drie accelerometers, drie gyroscopen en drie magnetometers. Sensorfusie werd gebruikt voor het verkrijgen van een meer betrouwbare schatting van bewegingen uitgevoerd door de deelnemers. Support Vector Machines (SVMs) werden toegepast om de prestaties van iedere patiënt op de genoemde taken te classificeren in één van de vijf mogelijke categorieën van de UPDRS (0 t/m 4). De score van de UPDRS van één van de vier beoordelaars diende als labels voor het classificatie algoritme. Door middel van Leave-One-Out Cross-validatie (LOOCV) werd de SVM getraind en de prestatie van het classificatie algoritme werd vergeleken met de evaluatie door de andere beoordelaars. Vervolgens werd deze procedure herhaald met de scores van de andere beoordelaars als labels voor

het classificatie algoritme. Daarnaast wordt in dit hoofdstuk de prestatie van classificatie algoritmen die gebruik maken van kenmerken ('features') speciaal ontwikkeld voor dit probleem, vergeleken met classificatie algoritmen die gebruik maken van data gedreven kenmerken. Voor de SVM gebaseerd op specifieke kenmerken, was de beste prestatie voor tikken met de vinger een classificatie fout van 31.5%. Dit was 16.5 % lager dan de gemiddelde inter-rater fout. De beste prestatie voor diadochokinese was een classificatie fout van 36%, die 9.8% lager was dan de gemiddelde inter-rater fout. Tenslotte, voor het tikken met de teen, resulteerde de beste prestatie van de SVM in een classificatie fout van 35.5%, die 20.2% lager was dan de gemiddelde inter-rater fout.

In **Hoofdstuk IV** wordt vervolgd met de analyse van bradykinesie. Het doel van dit hoofdstuk was het verkrijgen van lineaire modellen voor de score van drie van de bradykinesie-gerelateerde items van de UPDRS. Achterwaartse lineaire regressie werd toegepast voor het verkrijgen van een continu model voor het scoren van tikken met de vinger, diadochokinese en tikken met de teen. We verwachten dat continue schalen het mogelijk zullen maken om kleine bewegingen of verslechtering van de toestand van de patiënt te identificeren. Vier beoordelaars evalueerden deze taken op basis van de UPDRS. De scores van de beoordelaars werden vervolgens gemiddeld om mogelijke effecten van outliers te verminderen. LOOCV werd toegepast en de test fout werd gedefinieerd als het verschil tussen de voorspelde score en de werkelijke score. Voor tikken met de vinger was de kruisvalidatie fout 0.38; dit was 0.19 lager dan het gemiddelde verschil over alle beoordelaars. Voor diadochokinese was de kruisvalidatie fout 0.56, en daarmee 0.03 lager dan het gemiddelde verschil over alle beoordelaars. Tenslotte, voor tikken met de teen was de kruisvalidatie fout 0.48, wat 0.22 lager was dan het gemiddelde.

Het doel van **Hoofdstuk V** was het definiëren van een automatisch classificatie algoritme dat kan helpen EOA en DCD te differentiëren, wat momenteel nog een moeilijke taak is voor artsen. Bij kinderen zijn zowel EOA als DCD aandoeningen waarbij de coördinatie is aangedaan. Deze studie was gericht op het analyseren van de vinger-naar-neus taak van de UPDRS. Kinderen gediagnosticeerd met EOA of DCD en gezonde leeftijdsgenoten namen deel aan het onderzoek. Drie IMUs, sensorfusie en een 3D model werden gebruikt om de bewegingstrajecten van de wijsvinger van de deelnemers te schatten terwijl zij de vinger-naar-neus taak, zoals beschreven in de SARA, uitvoerden. Voor dit probleem werden specifieke kenmerken van de beweging bepaald om in het classificatie algoritme te gebruiken. Het random forest classificatie algoritme werd gebruikt om iedere taakuitvoering toe te wijzen aan een van de drie

mogelijke categorieën (EOA, DCD of gezond). Om een betere schatting van prestaties van het classificatie algoritme te verkrijgen, is multiple cross validatie gebruikt en zijn de resultaten gemiddeld. De totale gemiddelde classificatie nauwkeurigheid van de cross validaties was $70 \pm 20\%$. Op groepsniveau verkregen we een nauwkeurigheid van $31 \pm 8\%$ voor DCD, $77.8 \pm 0\%$ voor EOA en $82.9 \pm 30\%$ voor de gezonde controlepersonen. De belangrijkste kenmerken voor classificatie waren gerelateerd aan de onregelmatigheid en kromming van het 3D traject.

Tot slot vervolgen we in **Hoofdstuk VI** de analyse van coördinatie in EOA, DCD en leeftijdsgematchte gezonde kinderen. Het doel van dit hoofdstuk was het definiëren van een classificatie algoritme gebaseerd op loopdata vastgelegd door drie IMUs die bevestigd waren op de scheenbenen en op de onderrug. Kenmerken van de beweging werden bepaald op basis van bestaande literatuur en geselecteerd voor het classificatie algoritme door een methode die gebruik maakt van sequentiële achterwaartse selectie gebaseerd op de nauwkeurigheid van de kruisvalidatie. Ook hebben we Hidden Markov Models (HMM) voor elke categorie gedefinieerd en de likelihood bepaald van elke taakuitvoering om tot één van de drie modellen te horen. Deze zijn ook als kenmerken aan het classificatie model toegevoegd. De geselecteerde kenmerken werden gebruikt door een SVM die iedere taakuitvoering aan één van de diagnostische categorieën toekent. Voor het onderscheiden van EOA, DCD en gezonde kinderen werden 41 van de 90 kenmerken gebruikt. Iedere deelnemer liep meerdere keren. Voor deze individuele keren behaalde de classificatie een nauwkeurigheid van 63.8%. Na een meerderheid van stemmen strategie, waarbij iedere patiënt werd toegewezen aan die klasse waaraan de meeste van de keren werd toegewezen, werd een nauwkeurigheid van 78.4% behaald. Dit was 5.4% hoger dan de fenotypische evaluatie nauwkeurigheid van twee kinderneurologen.

Discussie

In dit proefschrift wordt een methode beschreven waarmee symptomen van patiënten met bewegingsstoornissen objectief kunnen worden beoordeeld. Hiervoor maakten we gebruik van gebruiksvriendelijke bewegingssensoren die vanwege hun kleine formaat en beperkte gewicht ook geschikt zijn voor toepassing in de klinische praktijk. Klinische toepassing kan leiden tot een verbetering van de zorg voor patiënten met bewegingsstoornissen. Zo helpt de techniek bij het correct diagnosticeren van complexe fenotypes, kan zij gebruikt

worden om het beloop van een bewegingsstoornis te monitoren, en is zij in staat om subtiele verbetering van klachten te signaleren.

In **Hoofdstuk II** kwantificeerden we de aanwezigheid van tremor. Op vergelijkbare wijze kwantificeerden we de ernst van bradykinesie, zoals in **Hoofdstuk III** en **IV** staat beschreven. Ten slotte kwantificeerden we in **Hoofdstukken V** en **VI** coördinatieproblemen om onderscheid te kunnen maken tussen vergelijkbare aandoeningen. Veel van de symptomen van bewegingsstoornissen die in dit proefschrift worden onderzocht, worden op dit moment visueel beoordeeld. In ieder hoofdstuk bespraken we een objectieve meetmethode die mogelijk van toegevoegde waarde kan zijn voor de diagnostiek en monitoring van verschillende bewegingsstoornissen. Vergeleken met de beoordelingsmethoden die momenteel in de kliniek worden gebruikt, presteerde onze objectieve analyse even goed of beter.

Wij denken dat het objectief beoordelen en continu meten van bewegingsstoornissen niet alleen zal leiden tot een verbetering van de klinische beoordeling, maar dat het ook onze kennis over de karakteristieke symptomen zal vergroten. Klinische toepassing van deze methode, waarbij de symptomen van een patiënt continu worden gemeten, levert namelijk veel gegevens op. Deze grote datasets kunnen vervolgens gebruikt worden voor het trainen van nog geavanceerdere classificatie algoritmes die ook goed zullen presteren bij de beoordeling van nieuwe patiënten. We verwachten dat het toenemende en continue gebruik van sensoren in de kliniek ervoor zal zorgen dat dit scenario inderdaad waarheid kan worden.

In dit proefschrift bespraken we een aantal technieken waarmee de symptomen van bewegingsstoornissen geanalyseerd en gekwantificeerd kunnen worden. Voor het ontwikkelen van de algoritmes gebruikten we echter relatief kleine datasets. Toekomstige beschikbaarheid van grotere datasets zal bijdragen aan de ontwikkeling van algoritmes die nog beter presteren. Daarnaast verwachten we dat de sensoren voor klinische bewegingsanalyse ook buiten het ziekenhuis zullen worden toegepast. Zo worden draagbare sensoren ('wearables') tegenwoordig veelvuldig gebruikt tijdens het sporten of voor het meten van de dagelijkse lichaamsbeweging. We denken dat deze trend zich ook zal voordoen bij sensoren voor klinische diagnostiek. Zo'n ontwikkeling zal artsen helpen om de symptomen van patiënten goed te monitoren en om deze ook buiten het ziekenhuis te kunnen beoordelen. Dat vergroot het inzicht in de factoren die van invloed zijn op bewegingsstoornissen, die nu meestal alleen in een klinische setting worden beoordeeld, waardoor de beoordeling van en de kennis over deze aandoeningen steeds verder zullen verbeteren.

Conclusie

In dit proefschrift gebruikten we bewegingssensoren om verschillende symptomen van bewegingsstoornissen objectief te beoordelen. Bij het ontwikkelen van deze beoordelingsmethode pasten we verschillende analysetechnieken toe. Uit onze onderzoeksresultaten blijkt dat het combineren van bewegingssensoren met 'machine learning'-technieken van toegevoegde waarde kan zijn bij het diagnosticeren en monitoren van bewegingsstoornissen. We verwachten dat de toenemende beschikbaarheid van bewegingssensoren, hun steeds kleinere formaat en lagere prijs zullen bijdragen aan de klinische implementatie van deze objectieve beoordelingsmethode. Daarnaast denken we dat daardoor grotere datasets beschikbaar zullen komen. Deze data kunnen vervolgens gebruikt worden voor het trainen van geavanceerdere classificatie algoritmes die uiteindelijk nog beter kunnen omgaan met nieuwe gegevens. De onderzoeken uit dit proefschrift zijn een voorbeeld van de actuele ontwikkeling binnen de geneeskunde waarbij er steeds meer belang wordt gehecht aan het verkrijgen van meer gegevens over patiënten. We denken deze ontwikkeling ertoe zal leiden dat het kwantificeren en continu monitoren van symptomen een belangrijke rol zal gaan spelen bij de klinische beoordeling van bewegingsstoornissen.

